



# UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE  
United States Patent and Trademark Office  
Address: COMMISSIONER FOR PATENTS  
P.O. Box 1450  
Alexandria, Virginia 22313-1450  
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
10/661,245	09/12/2003	Bhashyam Ramesh	NCR 11092	8704

7590 03/28/2005

John D. Cowart  
Teradata Law IP, WHQ-4W  
NCR Corporation  
1700 S. Patterson Blvd.  
Dayton, OH 45479-0001

EXAMINER

CORRIELUS, JEAN M

ART UNIT PAPER NUMBER

2162

DATE MAILED: 03/28/2005

Please find below and/or attached an Office communication concerning this application or proceeding.

## Office Action Summary

Application No.

10/661,245

Applicant(s)

RAMESH ET AL.

Examiner

Jean M Corrielus

Art Unit

2162

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --

### Period for Reply

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE \_\_\_\_ MONTH(S) FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If the period for reply specified above is less than thirty (30) days, a reply within the statutory minimum of thirty (30) days will be considered timely.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

### Status

- 1) ☐ Responsive to communication(s) filed on \_\_\_\_.
- 2a) ☐ This action is FINAL. 2b) ☒ This action is non-final.
- 3) ☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

### Disposition of Claims

- 4) ☐ Claim(s) \_\_\_\_ is/are pending in the application.
- 4a) Of the above claim(s) \_\_\_\_ is/are withdrawn from consideration.
- 5) ☐ Claim(s) \_\_\_\_ is/are allowed.
- 6) ☐ Claim(s) \_\_\_\_ is/are rejected.
- 7) ☐ Claim(s) \_\_\_\_ is/are objected to.
- 8) ☐ Claim(s) \_\_\_\_ are subject to restriction and/or election requirement.

### Application Papers

- 9) ☐ The specification is objected to by the Examiner.
- 10) ☐ The drawing(s) filed on \_\_\_\_ is/are: a) ☐ accepted or b) ☐ objected to by the Examiner.  
Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).  
Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).
- 11) ☐ The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

### Priority under 35 U.S.C. § 119

- 12) ☐ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).
- a) ☐ All b) ☐ Some \* c) ☐ None of:
1. ☐ Certified copies of the priority documents have been received.
2. ☐ Certified copies of the priority documents have been received in Application No. \_\_\_\_.
3. ☐ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

\* See the attached detailed Office action for a list of the certified copies not received.

### Attachment(s)

- 1) ☒ Notice of References Cited (PTO-892)
- 2) ☐ Notice of Draftsperson's Patent Drawing Review (PTO-948)
- 3) ☒ Information Disclosure Statement(s) (PTO-1449 or PTO/SB/08)  
Paper No(s)/Mail Date \_\_\_\_.
- 4) ☐ Interview Summary (PTO-413)  
Paper No(s)/Mail Date. \_\_\_\_.
- 5) ☐ Notice of Informal Patent Application (PTO-152)
- 6) ☐ Other: \_\_\_\_.

### DETAILED ACTION

1. This office action is in response to the request for consideration filed on November 30, 2004, which claims 1-12 are presented for further examination.

#### *Response to Arguments*

2. Applicant's arguments filed November 30, 2004 have been fully considered but they are not persuasive. (See below).

#### *Information Disclosure Statement*

3. The information disclosure statement (IDS) filed on November 30, 2004 complies with the provisions of M.P.E.P. 609. It has been placed in the application file. The information referred to therein has been considered as to the merits.

#### *Drawings*

The drawings are objected to under 37 CFR 1.83(a). The drawings must show every feature of the invention specified in the claims. Therefore, the **“identifying R unique n-gram T1...R in the string; for every unique n-gram Ts: if the frequency of Ts in a set of n-gram statistics is not greater than a first threshold: associating the string with a cluster associated with Ts; otherwise: for every other n-gram Tv in the string T1...R, except s: if the frequency of n-gram Tv is greater than the first threshold: if the frequency of n-gram pair Ts-Tv is not greater than a second threshold: associating the string with a cluster associated with the n-gram pair Ts-Tv; otherwise: for every other n-gram Tx in the string T1...R, except s and v: associating the string with a cluster associated with the n-gram triple Ts-Tv-Tx “ and “identifying R unique n-grams T1...R in the string; for every unique n-gram Ts: if the frequency of Ts in a set of n-gram statistics is not greater than a first threshold: associating the string with**

Art Unit: 2162

*a cluster associated with  $T_s$ ; otherwise: for  $i = 1$  to  $Y$ : for every unique set of  $i$   $n$ -grams  $T_u$  in the string  $T_1...R$ , except  $s$ : if the frequency of the  $n$ -gram set  $T_s-T_u$  is not greater than a second threshold: associating the string with a cluster associated with the  $n$ -gram set  $T_s-T_u$ ; if the string has not been associated with a cluster with this value of  $T_s$ : for every unique set of  $Y+1$   $n$ -grams  $T_{uy}$  in the string  $T_1...R$ , except  $s$ : associating the string with a cluster associated with the  $Y+2$   $n$ -gram group  $T_s-T_{uy}$*  must be shown or the feature(s) canceled from the claim(s). No new matter should be entered.

Corrected drawing sheets in compliance with 37 CFR 1.121(d) are required in reply to the Office action to avoid abandonment of the application. Any amended replacement drawing sheet should include all of the figures appearing on the immediate prior version of the sheet, even if only one figure is being amended. The figure or figure number of an amended drawing should not be labeled as "amended." If a drawing figure is to be canceled, the appropriate figure must be removed from the replacement sheet, and where necessary, the remaining figures must be renumbered and appropriate changes made to the brief description of the several views of the drawings for consistency. Additional replacement sheets may be necessary to show the renumbering of the remaining figures. Each drawing sheet submitted after the filing date of an application must be labeled in the top margin as either "Replacement Sheet" or "New Sheet" pursuant to 37 CFR 1.121(d). If the changes are not accepted by the examiner, the applicant will be notified and informed of any required corrective action in the next Office action. The objection to the drawings will not be held in abeyance.

***Remark***

4. Applicants asserted that the claimed subject matter is described in the specification page 6, line 19 through page 8, line 15. The examiner disagrees with the precedent assertion.

However, the passage of the specification as indicated by the Applicants does not detail the invention as claimed. Applicants are advised to show where in the specification each limitation of the claimed language. For Applicants clarification the examiner has reinstated the rejection under 35 U.S.C. 112 and 101 below.

***Claim Rejections - 35 USC § 112***

5. The following is a quotation of the first paragraph of 35 U.S.C. 112:

The specification shall contain a written description of the invention, and of the manner and process of making and using it, in such full, clear, concise, and exact terms as to enable any person skilled in the art to which it pertains, or with which it is most nearly connected, to make and use the same and shall set forth the best mode contemplated by the inventor of carrying out his invention.

6. Claims 1-3 and 6-12 are rejected under 35 U.S.C. 112, first paragraph, as failing to comply with the written description requirement. The claim(s) contains subject matter which was not described in the specification in such a way as to reasonably convey to one skilled in the relevant art that the inventor(s), at the time the application was filed, had possession of the claimed invention. Claims 1 and 10 recite “identifying R unique n-gram T1...R in the string; for every unique n-gram Ts: if the frequency of Ts in a set of n-gram statistics is not greater than a first threshold: associating the string with a cluster associated with Ts; otherwise: for every other n-gram Tv in the string T1...R, except s: if the frequency of n-gram Tv is greater than the first threshold: if the frequency of n-gram pair Ts-Tv is not greater than a second threshold: associating the string with a cluster associated with the n-gram pair Ts-Tv;

Art Unit: 2162

otherwise: for every other n-gram  $T_x$  in the string  $T_1...R$ , except  $s$  and  $v$ : associating the string with a cluster associated with the n-gram triple  $T_s-T_v-T_x$ "; and

*claim 6 recites "identifying  $R$  unique n-grams  $T_1...R$  in the string; for every unique n-gram  $T_s$ : if the frequency of  $T_s$  in a set of n-gram statistics is not greater than a first threshold: associating the string with a cluster associated with  $T_s$ ; otherwise: for  $i = 1$  to  $Y$ : for every unique set of  $i$  n-grams  $T_u$  in the string  $T_1...R$ , except  $s$ : if the frequency of the n-gram set  $T_s-T_u$  is not greater than a second threshold: associating the string with a cluster associated with the n-gram set  $T_s-T_u$ ; if the string has not been associated with a cluster with this value of  $T_s$ : for every unique set of  $Y+1$  n-grams  $T_{uy}$  in the string  $T_1...R$ , except  $s$ : associating the string with a cluster associated with the  $Y+2$  n-gram group  $T_s-T_{uy}$ ". The specification page 6, line 19 through pages 8, line 15 as indicated by the Applicants does not provide any detail of the above-mentioned limitations of the claim. The examiner has read the claims using the specification, however, the claimed limitations can be found into the specification. Applicants are advised to point out wherein the specification the limitations of the claims are detailed. Applicants have the opportunity to amend the specification or cancel the limitations from the claims. Applicants are reminded that no new matter should be added.*

Art Unit: 2162

7. The following is a quotation of the second paragraph of 35 U.S.C. 112:

The specification shall conclude with one or more claims particularly pointing out and distinctly claiming the subject matter which the applicant regards as his invention.

8. Claims 4-5 are rejected under 35 U.S.C. 112, second paragraph, as being indefinite for failing to particularly point out and distinctly claim the subject matter which applicant regards as the invention. Claim 4, lines 5 and line 7 recites "if any". Such language provides uncertainty or doubt, as whether the steps of associating each string with clusters will achieve. "if any" does not guarantee a completion of the associating step rather than a possibility of associating each string with clusters associated with low frequency n-grams from that string; and associating each string with clusters associated with low frequency pairs of high frequency n-grams from that string if it is existed. Applicant is advised to amend the claims to clarify that uncertainty set forth in the claims.

***Claim Rejections - 35 USC § 101***

9. 35 U.S.C. 101 reads as follows:

Whoever invents or discovers any new and useful process, machine, manufacture, or composition of matter, or any new and useful improvement thereof, may obtain a patent therefor, subject to the conditions and requirements of this title.

10. Claims 1-3 and 6-9 are rejected under 35 U.S.C. 101 because the claimed invention is directed to non-statutory subject matter.

Claims 1-3 and 6-9, in view of **MPEP section 2106 IV.B.2.(b)** are not statutory because they merely recite a number of computing steps without producing any tangible result and/or being limited to a practical application within the technological arts. The language of the claim raises a question as to whether the claim is directed merely to an abstract idea that is not tied to a

Art Unit: 2162

technological art, environment or machine which would result in a practical application producing a concrete, useful, and tangible result to form the basis of statutory subject matter under 35 U.S.C. 101.

With regarding claims 1, 4 and 6:

While the preamble of the claim states, “a method for clustering a string including a plurality of characters”, the claim fails to contain a computer that is used implemented the method for clustering a string so as to realize its functionality. Thus, claim 1 is merely abstract idea whereby “clustering a string including a plurality of characters” is being processed without any links to a practical result in the technology arts and without computer manipulation.

With regarding claims 2-3, 5 and 7-9:

The dependent claims 2-3, 5 and 7-9 are rejected for fully incorporating the errors of their respective base claims by dependency. Thus, claim 2-3, 5 and 7-9 are merely abstract idea and are being processed without any links to a practical result in the technology arts and without computer manipulation.

### ***Claim Rejections - 35 USC § 103***

11. The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all obviousness rejections set forth in this Office action:

(a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negated by the manner in which the invention was made.

12. This application currently names joint inventors. In considering patentability of the claims under 35 U.S.C. 103(a), the examiner presumes that the subject matter of the various



Art Unit: 2162

claims was commonly owned at the time any inventions covered therein were made absent any evidence to the contrary. Applicant is advised of the obligation under 37 CFR 1.56 to point out the inventor and invention dates of each claim that was not commonly owned at the time a later invention was made in order for the examiner to consider the applicability of 35 U.S.C. 103(c) and potential 35 U.S.C. 102(e), (f) or (g) prior art under 35 U.S.C. 103(a).

13. Claims 1-12 best understood by the examiner are rejected under 35 U.S.C. 103(a) as being unpatentable over Kreulen et al., (hereinafter "Kreulen") US Patent No. 6,862,586 and Chandrasekar et al., (hereinafter "Chandrasekar") US Patent no. 6,578,032.

Claims 1, 6 and 10 can only be interpreted as best understood by the examiner.

As to claims 1 and 10, Chandrasekar discloses "A method for clustering a plurality of strings, each string including a plurality of characters" as a use of providing a method for clustering character strings (col.2, lines 23-25). In particular, Chandrasekar discloses the claimed "identifying R unique n-grams T1...R in the string" (col.2, lines 3-10; col.7, lines 30-45); "if the frequency of Ts in a set of n-gram statistics is not greater than a first threshold: associating the string with a cluster associated with Ts" (col.12, line 59-col.12, line 14); "for every other n-gram Tv in the string T1...R, except s: if the frequency of n-gram Tv is greater than the first threshold: if the frequency of n-gram pair Ts-Tv is not greater than a second threshold: associating the string with a cluster associated with the n-gram pair Ts-Tv" (col.12, line 59-col.12, line 14). However, Chandrasekar does not explicitly disclose the use wherein "for every other n-gram Tx in the string T1...R, except s and v: associating the string with a cluster associated with the n-gram triple Ts-Tv-Tx". On the other hand, Kreulen discloses a method of searching a database using query, clustering the result items into logical categories and ranking the each categories

Art Unit: 2162

based on the frequency of the occurrence of words (col.1, line 67-col.2, line 3). In particular, Kreulen discloses the claimed “for every other n-gram Tx in the string T1...R, except s and v: associating the string with a cluster associated with the n-gram triple Ts-Tv-Tx” (col.4, lines 50-56). Therefore, it would have been obvious to one having ordinary skill in the art at the time the invention was made to combine the teachings of the cited references. One having ordinary skill in the art would have found it motivated to create an automated grouping using a clustering technique in order to provide easy update with the advent of a computer system.

As to claim 6, Chandrasekar discloses “A method for clustering a plurality of strings, each string including a plurality of characters” as a use of providing a method for clustering character strings (col.2, lines 23-25). In particular, Chandrasekar discloses the claimed “identifying R unique n-grams T1...R in the string” (col.2, lines 3-10; col.7, lines 30-45); “if the frequency of Ts in a set of n-gram statistics is not greater than a first threshold” (col.12, line 59-col.12, line 14); “associating the string with a cluster associated with Ts; otherwise: for i = 1 to Y: for every unique set of i n-grams Tu in the string T1...R, except s: if the frequency of the n-gram set Ts-Tu is not greater than a second threshold: associating the string with a cluster associated with the n-gram set Ts-Tu” (col.12, line 59-col.12, line 14). However, Chandrasekar does not explicitly disclose the use wherein “; if the string has not been associated with a cluster with this value of Ts: for every unique set of Y+1 n-grams Tuy in the string T1...R, except s: associating the string with a cluster associated with the Y+2 n-gram group Ts-Tuy”. On the other hand, Kreulen discloses a method of searching a database using query, clustering the result items into logical categories and ranking the each categories based on the frequency of the occurrence of words

Art Unit: 2162

(col.1, line 67-col.2, line 3). In particular, Kreulen discloses the claimed “if the string has not been associated with a cluster with this value of  $T_s$ : for every unique set of  $Y+1$  n-grams  $T_{uy}$  in the string  $T_1...R$ , except  $s$ : associating the string with a cluster associated with the  $Y+2$  n-gram group  $T_s-T_{uy}$ ” (col.4, lines 50-56). Therefore, it would have been obvious to one having ordinary skill in the art at the time the invention was made to combine the teachings of the cited references. One having ordinary skill in the art would have found it motivated to create an automated grouping using a clustering technique in order to provide easy update with the advent of a computer system.

As to claim 7, Kreulen discloses “ $Y=1$ ” (col.5, lines 10-col.6).

As to claims 2-3 and 8-9, Kreulen discloses the claimed “compiling n-gram statistics, and pair statistic and group statistics” (col.2, lines 36-67).

As to claim 4, Chandrasekar discloses “A method for clustering a plurality of strings, each string including a plurality of characters” as a use of providing a method for clustering character strings (col.2, lines 23-25). In particular, Chandrasekar discloses the claimed “identifying unique n-grams in each string” (col.2, lines 3-10; col.7, lines 30-45). However, Chandrasekar does not explicitly disclose the use of associating each string with clusters associated with low frequency n-grams from that string, if any and associating each string with clusters associated with low-frequency pairs of high frequency n-grams from that string, if any. On the other hand, Kreulen discloses a method of searching a database using query, clustering the result items into logical

Art Unit: 2162

categories and ranking the each categories based on the frequency of the occurrence of words (col.1, line 67-col.2, line 3). In particular, Kreulen discloses the claimed “associating each string with clusters associated with low frequency n-grams from that string, if any” (col.4, lines 50-56); and “associating each string with clusters associated with low-frequency pairs of high frequency n-grams from that string, if any” (col.4, lines 50-56). Therefore, it would have been obvious to one having ordinary skill in the art at the time the invention was made to combine the teachings of the cited references, wherein the Editorial database, provided therein (see Chandrasekar’s fig.8) would incorporate the use of associating each string with clusters associated with low frequency n-grams from that string, if any and associating each string with clusters associated with low-frequency pairs of high frequency n-grams from that string, if any, in the same conventional manner as disclosed by Kreulen(col.4, lines 50-56). One having ordinary skill in the art would have found it motivated to create an automated grouping using a clustering technique in order to provide easy update with the advent of a computer system.

As to claims 5, Kreulen discloses the claimed “ where a string does not include any low-frequency pairs of high frequency n-grams associating that string with clusters associated with triples of n-grams including the pair” (col.3, lines 13-16; col.4, lines 57-61; col.5, lines 50-53; col.6, lines 52-55).

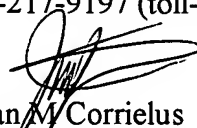
Art Unit: 2162

*Conclusion*

14. Any inquiry concerning this communication or earlier communications from the examiner should be directed to Jean M Corrielus whose telephone number is (571) 272-4032. The examiner can normally be reached on 10 hours shift.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, John Breene can be reached on (571) 272-4107. The fax phone number for the organization where this application or proceeding is assigned is 703-872-9306.

Information regarding the status of an application may be obtained from the Patent Application Information Retrieval (PAIR) system. Status information for published applications may be obtained from either Private PAIR or Public PAIR. Status information for unpublished applications is available through Private PAIR only. For more information about the PAIR system, see <http://pair-direct.uspto.gov>. Should you have questions on access to the Private PAIR system, contact the Electronic Business Center (EBC) at 866-217-9197 (toll-free).

  
Jean M Corrielus  
Primary Examiner  
Art Unit 2162

March 20, 2005